First insights into the genetic diversity of *Mycobacterium tuberculosis* strains in Salvador, Bahia State, Brazil

Primeiro ensaio sobre diversidade genética das cepas de *Mycobacterium tuberculosis* em Salvador, Bahia, Brasil

Aída Cristina do Nascimento Silva [1]
Lucilaine Ferrazoli [2]
Vera Simonsen [2]
Joice Neves Reis [3]
Susan Martins Pereira [4]
Theomira Mauadie de Azevedo Carmo [1]
Eduardo Luiz Andrade Mota [4]
Mitermayer Galvão Reis [1]

[1] *Centro de Pesquisa Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brasil.*
[2] *Instituto Adolfo Lutz, São Paulo, Brasil.*
[3] *Faculdade de Farmácia, Universidade Federal da Bahia, Salvador, Brasil.*
[4] *Instituto de Saúde Coletiva, Universidade Federal da Bahia, Salvador, Brasil.*

Correspondence
A. C. N. Silva
Centro de Pesquisa Gonçalo Moniz, Fundação Oswaldo Cruz.
Rua Waldemar Falcão 121, Salvador, BA 40296-710, Brasil.
aida@aluno.bahia.fiocruz.br

Abstract

This study constitutes a first attempt to describe the genetic population structure of Mycobacterium tuberculosis *circulating in Salvador, Bahia State, Brazil. A total of 56 confirmed cases of pulmonary tuberculosis, identified between March and June 2008, were analyzed using restriction fragment length polymorphism (IS6110-RFLP). The study population was characterized by a predominance of males (71.43%) over 30 years of age (68.75%). Forty-one isolates were found to belong to a single pattern (73.2%), while 15 (26.7%) were found in group patterns, forming six clusters. The higher level of diversity observed is much more suggestive of endogenous reactivation than recent transmission.*

Tuberculosis; Molecular Epidemiology; Genetic Variation; Mycobacterium tuberculosis

Introduction

Tuberculosis (TB) is a global public health issue, with 9.27 million new cases and 1.3 million deaths worldwide in 2007 [1]. Brazil ranks among the 22 countries with the highest incidence rates of tuberculosis, with 100,000 new cases of TB being registered annually. In 2006, 6,500 cases of TB were registered in the State of Bahia, which is currently ranked as the third state in absolute numbers of cases in Brazil. Of these cases, 2,300 occurred in the city of Salvador [2]. Recently, molecular techniques have been used to explain the epidemiological aspects of TB, contributing towards improving understanding of transmission dynamics that drive the disease [3].

To date, no studies have been carried out on the epidemiologic and genetic diversity of circulating *Mycobacterium tuberculosis* in Salvador, a city in the Northeast Region of Brazil, where the incidence of this disease is high. The aim of the present study was to describe the genetic population structure of *M. tuberculosis* circulating in Salvador.

Methodology

Patients

This study included confirmed TB cultures from patients over 18 years of age diagnosed between March and June 2008 and receiving care at the

Otavio Mangabeira Specialist Hospital (HEOM/SESAB), a referral unit for pneumopathies situated in the city of Salvador. The research project was approved by the National Health Surveillance Agency/Ministry of Health in Brasília, and by the ethics committee of the Collective Health Institute, Federal University of Bahia.

Bacterial isolates

The *M. tuberculosis* isolates were collected and processed from sputum samples in accordance with the protocol established by the Department of Mycobacteriology of the Central Laboratory of Public Health "Professor Gonçalo Muniz" (LACEN-BA) [4].

IS6110-RFLP analysis and definitions

The clinical isolates of *M. tuberculosis* were typed using restriction fragment length polymorphism (IS6110-RFLP) analysis in accordance with the standardized protocol described by van Embden et al. [5]. The genetic profile of each strain was analyzed and visualized with GelCompar II version 4.0 (Applied Maths, Kortrijk, Belgium) using *M. tuberculosis* 14323 as the reference strain. All the strains with a similarity coefficient of > 90% were considered part of a "family". Within the molecular typing system, cluster strains were defined as being those which were found to have the same RFLP pattern (number of bands and position of the fragment) in two or more isolates of *M. tuberculosis*. Isolates which were found to have RFLP patterns distinctly different from the other patterns identified were considered to be non-clusters [3]. Isolates forming part of a cluster pattern were defined as representing a recently acquired infection and standard non-cluster isolates were defined as representing reactivation of an infection acquired in the past [3].

Data analysis

Exploratory data management was performed using Stata version 7.0 (Stata Corp., College Station, USA). Prevalence ratios (PR) and their respective 95% confidence intervals (95%CI) were calculated.

Findings

Analysis of patient characteristics and associated factors

Fifty six patients whose sputum cultures tested positive for *M. tuberculosis* were enrolled in this study between March and June 2008,. Of this total, the majority (71.43%) were male and over 30 years of age (68.75%) (Table 1). Regarding ethnic background, most of the patients (80%) classified themselves as non-white. With respect to education, 63.89% of the patients were illiterate or had failed to complete primary school. A quarter of the patients did not reside in Salvador and 23.4% had been submitted to a category III treatment regimen [indicated in Brazil for the treatment of cases in which previous treatment with regimens I, IR (recurrence after cure) and II, has failed].

A positive association was found between the formation of cluster patterns and the following factors: being male (PR = 1.87; 95%CI: 0.62-5.63), age > 30 years (PR = 1.67; 95%CI: 0.54-5.11), being self-classified as non-white (PR = 3.00; 95%CI: 0.46-19.35), being illiterate or having failed to complete primary school (PR = 1.70; 95%CI: 0.55-5.17); not living in Salvador (PR = 1.67; 95%CI: 0.69-4.00); having been submitted to a category III treatment regimen for tuberculosis (PR = 1.31; 95%CI: 0.51-3.36). The associations found were not statistically significant.

RFLP analysis

Of the 56 patients analyzed, 41 of the isolates (73.21%) had single patterns and 15 (26.78%) had group patterns. Overall, all the isolates had 2-16 copies of IS6110-RFLP and 10.7% of all samples analyzed had six or fewer copies of this repetitive element. Fifteen isolates formed 6 distinct clusters denominated CL I, CL II, CL III, CL IV, CL V and CL VI. Each cluster was composed of 2 isolates except for CL V, which was composed of 5 isolates (Figure 1).

Discussion

This is the first exploratory study conducted in Bahia with the aim of describing the genetic population structure of *M. tuberculosis*, as regards recent infection or reactivation. The principal finding refers to the fact that 26.78% of isolates formed cluster patterns, suggesting recent transmission of TB in this group of patients. This proportion is slightly higher than the findings of studies conducted in Rio de Janeiro using similar methodology (19%) [6] and similar to the results of studies carried out in isolates in Porto Alegre (State of Rio Grande do Sul) and São Paulo [3,7] which amounted to 29.1% and 33%, respectively. In the present study, the percentage of cluster strains found was slightly lower than those reported in studies conducted in France (35%) [8] and San Francisco (40%) [9].

Table 1

Bivariate analysis of the association between the co-variables and restriction fragment length polymorphism (RFLP) pattern (cluster or non-cluster) of *Mycobacterium tuberculosis* isolates from patients infected with tuberculosis (TB). Salvador, Bahia State, Brazil.

Predictive variables	Cluster (n = 15)		Non-cluster (n = 41)		Total (N = 56) *		PR (95%CI) **
	n	%	n	%	n	%	
Sex							
Male	14	35.00	26	65.00	40	71.43	1.87 (0.62-5.63)
Female	3	18.75	13	81.25	16	28.57	1.00
Age (years)							
> 30	11	33.33	22	66.67	33	68.75	1.67 (0.54-5.11)
18-29	3	20.00	12	80.00	15	31.25	1.00
Self-reported ethnic background							
Non-white	12	42.86	16	57.14	28	80.00	3.00 (0.46-19.35)
White	1	14.29	6	85.71	7	20.00	1.00
Education level ***							
Illiterate/Did not finish primary school	9	39.13	14	60.87	23	63.89	1.70 (0.56-5.17)
Completed primary school/Attended at least some high school	3	23.08	10	76.92	13	36.11	1.00
Residence in Salvador							
No	5	41.67	7	58.33	12	25.00	1.67 (0.69-4.00)
Yes	9	25.00	27	75.00	36	75.00	1.00
Treatment regimen							
Category III (therapeutic failures)	4	36.36	7	63.64	11	23.40	1.31 (0.51-3.36)
Category I (new cases of TB)	10	27.78	26	72.22	36	76.60	1.00

PR: prevalence ratio; 95%CI: 95% confidence interval.

Note: missing data – age (missing in 8 cases); self-reported ethnicity (missing in 21 cases); education level (missing in 20 cases); residence in Salvador (missing in 8 cases); treatment regimen (missing in 9 cases).

* Total varies in accordance with availability of data;

** PR and 95%CI were calculated after excluding missing data;

*** Classification adopted in the registers of the notification and investigation of cases of tuberculosis (Departamento de Informática do SUS. Sistema de Informações de Agravos de Notificação. http://www.datasus.gov.br).

The proportion of non-cluster strains was 73.21%, suggesting the need to improve TB control, promoting preventive actions against reactivation. In Brazil, recent studies [3,7] show similar findings. Around 70% of cases, result from old infections (non-cluster strains), showing that the occurrence of new cases in the country originates predominantly from endogenous reactivation.

In this study, patients were from a heterogeneous group of people concentrated in the capital city, the Greater Metropolitan Region of Salvador and other towns in the state, in addition to two patients living in another state (São Paulo). This finding shows the mobility of the population-genetic characteristics of tuberculosis between different regions, so providing a contribution to implementing adequate measures in cases of recent infection as well as reactivation of an old infection.

One of the limitations of this study is that, due to the small sample size, it was not possible to obtain a statistically significant association between the formation of cluster patterns and the factors analyzed. Another limitation refers to the fact that the findings on the proportion of cluster patterns may be affected by the incidence of TB, by the heterogeneous nature of the population and by the sampling fraction included in the study. An analysis of these aspects would contribute towards obtaining a more accurate calculation of recent infection in a given community [6,10].

Despite these limitations, a wide genetic diversity was found in the strains of *M. tuberculosis* analyzed in TB patients in this referral hospital in the city of Salvador. Further studies should be carried out to acquire more data on this subject in Salvador and enable the investigation into associations between the various co-variables and the proportion of cluster strains in this city.
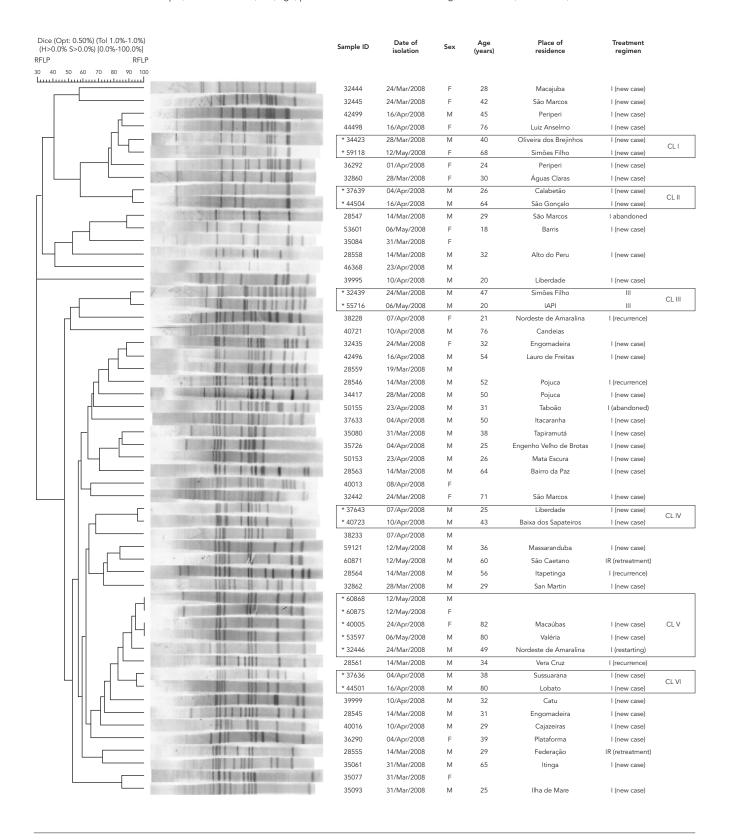
Figure 1

Dendrogram representing the different RFLP patterns found in isolates of *M. tuberculosis* from patients receiving care at a referral hospital, according to the identification number of the sample, date of isolation, sex, age, place of residence and treatment regimen. Salvador, Bahia State, Brazil.

Sample ID	Date of isolation	Sex	Age (years)	Place of residence	Treatment regimen	
32444	24/Mar/2008	F	28	Macajuba	I (new case)	
32445	24/Mar/2008	F	42	São Marcos	I (new case)	
42499	16/Apr/2008	M	45	Periperi	I (new case)	
44498	16/Apr/2008	F	76	Luiz Anselmo	I (new case)	
* 34423	28/Mar/2008	M	40	Oliveira dos Brejinhos	I (new case)	CL I
* 59118	12/May/2008	F	68	Simões Filho	I (new case)	
36292	01/Apr/2008	F	24	Periperi	I (new case)	
32860	28/Mar/2008	F	30	Águas Claras	I (new case)	
* 37639	04/Apr/2008	M	26	Calabetão	I (new case)	CL II
* 44504	16/Apr/2008	M	64	São Gonçalo	I (new case)	
28547	14/Mar/2008	M	29	São Marcos	I abandoned	
53601	06/May/2008	F	18	Barris		
35084	31/Mar/2008	F				
28558	14/Mar/2008	M	32	Alto do Peru	I (new case)	
46368	23/Apr/2008	M				
39995	10/Apr/2008	M	20	Liberdade	I (new case)	
* 32439	24/Mar/2008	M	47	Simões Filho	III	CL III
* 55716	06/May/2008	M	20	IAPI	III	
38228	07/Apr/2008	F	21	Nordeste de Amaralina	I (recurrence)	
40721	10/Apr/2008	M	76	Candeias		
32435	24/Mar/2008	F	32	Engomadeira	I (new case)	
42496	16/Apr/2008	M	54	Lauro de Freitas	I (new case)	
28559	19/Mar/2008	M				
28546	14/Mar/2008	M	52	Pojuca	I (recurrence)	
34417	28/Mar/2008	M	50	Pojuca	I (new case)	
50155	23/Apr/2008	M	31	Taboão	I (abandoned)	
37633	04/Apr/2008	M	50	Itacaranha	I (new case)	
35080	31/Mar/2008	M	38	Tapiramutá	I (new case)	
35726	04/Apr/2008	M	25	Engenho Velho de Brotas	I (new case)	
50153	23/Apr/2008	M	26	Mata Escura	I (new case)	
28563	14/Mar/2008	M	64	Bairro da Paz	I (new case)	
40013	08/Apr/2008	F				
32442	24/Mar/2008	F	71	São Marcos	I (new case)	
* 37643	07/Apr/2008	M	25	Liberdade	I (new case)	CL IV
* 40723	10/Apr/2008	M	43	Baixa dos Sapateiros	I (new case)	
38233	07/Apr/2008	M				
59121	12/May/2008	M	36	Massaranduba	I (new case)	
60871	12/May/2008	M	60	São Caetano	IR (retreatment)	
28564	14/Mar/2008	M	56	Itapetinga	I (recurrence)	
32862	28/Mar/2008	M	29	San Martin	I (new case)	
* 60868	12/May/2008	M				
* 60875	12/May/2008	F				
* 40005	24/Apr/2008	F	82	Macaúbas	I (new case)	CL V
* 53597	06/May/2008	M	80	Valéria	I (new case)	
* 32446	24/Mar/2008	M	49	Nordeste de Amaralina	I (restarting)	
28561	14/Mar/2008	M	34	Vera Cruz	I (recurrence)	
* 37636	04/Apr/2008	M	38	Sussuarana	I (new case)	CL VI
* 44501	16/Apr/2008	M	80	Lobato	I (new case)	
39999	10/Apr/2008	M	32	Catu	I (new case)	
28545	14/Mar/2008	M	31	Engomadeira	I (new case)	
40016	10/Apr/2008	M	29	Cajazeiras	I (new case)	
36290	04/Apr/2008	F	39	Plataforma	I (new case)	
28555	14/Mar/2008	M	29	Federação	IR (retreatment)	
35061	31/Mar/2008	M	65	Itinga	I (new case)	
35077	31/Mar/2008	F				
35093	31/Mar/2008	M	25	Ilha de Mare	I (new case)	

Resumo

Este é o primeiro estudo realizado na Bahia, Brasil, visando à descrição da estrutura da população genética circulante do Mycobacterium tuberculosis *na cidade de Salvador. Um total de 56 casos confirmados de tuberculose pulmonar, identificados entre março e junho de 2008, foi analisado pelo método* Restriction Fragment Lenght Polymorphism *(IS6110-RFLP). A população de estudo foi caracterizada como a maioria do sexo masculino (71,43 %), idade acima de 30 anos (68,75%). Quarenta e um isolados (73,21%) com padrão único, enquanto 15 (26,75%) apresentaram padrões agrupáveis, formando seis* clusters. *A alta taxa de diversidade das cepas de* M. tuberculosis *observada é mais sugestiva de reativação endógena do que transmissão recente.*

Tuberculose; Epidemiologia Molecular; Variação Genética; Mycobacterium tuberculosis

Contributors

Acknowledgments

References

1. World Health Organization. WHO report 2002: global tuberculosis control. http://www.who.in to/gtb/publications/index.htm (accessed on 02/ Dec/2007).

2. Secretaria de Vigilância em Saúde, Ministério da Saúde. Situação da tuberculose no Brasil. http:// www.saude.gov.br/svs/tuberculose (accessed on 21/Nov/2007).

3. Ferrazoli L, Palaci M, Marques LRM, Jamal LF, Afiune JB, Chimara E, et al. Transmission of tuberculosis in an endemic urban setting in Brazil. Int J Tuberc Lung Dis 2000; 4:18-25.

4. Coordenação Geral de Laboratórios de Saúde Pública, Departamento de Vigilância Epidemiológica, Centro de Referência Prof. Hélio Fraga, Secretaria de Vigilância em Saúde, Ministério da Saúde. Cultura. In: Coordenação Geral de Laboratórios de Saúde Pública, Departamento de Vigilância Epidemiológica, Centro de Referência Prof. Hélio Fraga, Secretaria de Vigilância em Saúde, Ministério da Saúde. Manual de bacteriologia da tuberculose. 3ª Ed. Rio de Janeiro: Centro de Referência Prof. Hélio Fraga; 2005. p. 59-74.

5. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol 1993; 31:406-9.

6. Fandinho FC, Kritski AL, Hofer C, Conde Junior H, Ferreira RM, Saad MH, et al. RFLP patterns and risk factors for recent tuberculosis transmission among hospitalized tuberculosis patients in Rio de Janeiro, Brazil. Trans R Soc Trop Med Hyg 2000; 94:271-5.

7. Borges M, Cafrune PI, Possuelo LG, Valim ARM, Ribeiro MO, Rossetti MLR. Análise molecular de cepas de *Mycobacterium tuberculosis* provenientes de um centro de saúde ambulatorial em Porto Alegre. J Bras Pneumol 2004; 30:448-54.

8. Gutiérrez MC, Vincent V, Aubert D, Bizet J, Gaillot O, Lebrun L. Molecular fingerprinting of *Mycobacterium tuberculosis* and risk factors for tuberculosis transmission in Paris, France, and surrounding area. J Clin Microbiol 1998; 36:486-92.

9. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, et al. The epidemiology of tuberculosis in San Francisco. Apopulation-based study using conventional and molecular methods. N Engl J Med 1994; 330:1703-9.

10. Glynn JR, Vynnycky E, Fine PEM. Influence of sampling on estimates of clustering and recent transmission of *Mycobacterium tuberculosis* derived from DNA fingerprinting techniques. Am J Epidemiol 1999; 149:366-71.